UNITED STATES DEPARTMENT OF COMMERCE
Bureau of the Census
Washington, DC 20233-0001

February 1, 2000

DSSD CENSUS 2000 PROCEDURES AND OPERATIONS MEMORANDUM SERIES R-24

MEMORANDUM FOR   Maureen Lynch
Assistant Division Chief, Computer Match Processing
Decennial Statistical Studies Division

From: Donna Kostanich $\text{\scriptsize DK}$
Assistant Division Chief, Sampling and Estimation
Decennial Statistical Studies Division

Prepared by: Matt Salganik $\text{\scriptsize MS}$
Decennial Statistical Studies Division

Subject: Accuracy and Coverage Evaluation Survey: Small Block Cluster
Subsampling

I.   Introduction

This memorandum provides instructions for the small block cluster subsampling operation.
Before this operation the Accuracy and Coverage Evaluation (A.C.E.) reduction sample will
contain 5,000 small clusters in the United States and 96 small clusters in Puerto Rico. Small
clusters are expected to have between zero and two housing units based on an early census
address list. Conducting interviewing and follow-up operations in clusters of this size is not as
cost-effective as in larger clusters. Therefore, to allocate A.C.E. resources more efficiently, we
will only include a subsample of these small clusters in the A.C.E. interviewing sample. The
same subsampling procedure will be used for clusters in the United States and Puerto Rico.

This subsampling operation will reduce the sample of small clusters while at the same time
attempting to balance among three goals. First, we would like to prevent any small clusters from
having weights that are extremely high compared to other clusters in the sample. Second, we
would like to have lower weights on clusters where the number of housing units is different than
we expected. These first two goals attempt to reduce the contribution of small clusters to the
variance of the dual system estimates. The third goal is to ensure that the Field Division can
efficiently manage the resulting workloads.

To achieve these goals we will use differential subsampling where the subsampling rates are
based on the number of keyed and valid housing units from the A.C.E. Independent Listing[1] (IL)
and the number of housing units on the Decennial Master Address File (DMAF). This DMAF

---

[1] This IL housing unit count includes units with the status 'future new construction.'

housing unit count will be based on the January 2000 update. All American Indian County[2] (AIC), American Indian Reservation (AIR), and List/Enumerate clusters will be retained to avoid increasing the weights on these clusters.

The exact subsampling rates have not yet been determined. See Attachment A for the approximate take-everys. The exact rates will be determined after the keyed and valid IL counts and January DMAF are available. Once the rates have been determined, they will be keyed into a parameter file which will be provided to you in late January of 2000, a few days after we are provided with the keyed and valid housing units counts.

Small block cluster subsampling is part of the larger process of selecting the sample clusters for the A.C.E. This process begins with the listing sample selection which yields a sample of approximately 2 million housing units. An independent listing operation is done to create an address list. Next, the listing sample, which was based on the design of the Integrated Coverage Measurement Survey, will be subsampled to yield the A.C.E. reduced sample, which will be based on the A.C.E. design. Small block cluster subsampling will then occur resulting in the A.C.E. sample clusters. These are the clusters that will be in the A.C.E. interview sample. However, there is one more sampling process before we arrive at our final 300,000 housing unit sample – large block cluster subsampling. In this process some housing units will be removed from sample in large block clusters (those with 80+ housing units). The remaining housing units after large block cluster subsampling make up the A.C.E. interview sample.

This specification should be used to flowchart the process, to generate further discussion on requirements, to identify and finalize record layouts of input and output files, and to write computer software to implement the methodology. During and after a testing phase, it is likely that changes to the specification will be necessary.

Any comments or questions should be directed to Matt Salganik (301-457-3636) or Debbie Fenstermaker (301-457-4195).

II. Assumptions

   A. The A.C.E. block cluster reduction has been completed.

   B. All independent listing counts used will be 'keyed and valid' counts.

   C. For purposes of small block cluster subsampling, the independent list housing unit counts include units coded as 'future new construction'.

---

[2]American Indian Country includes Tribal Jurisdiction Statistical Areas, Tribal Designated Statistical Areas, Alaska Native Village Statistical Areas, and American Indian Reservations and associated trustlands. Throughout this specification the term American Indian Country will exclude American Indian Reservations and associated trustlands.

D.   The definition of American Indian Country is as of spring 1999 when block clustering was completed. This information could change before all census operations are completed.

E.   Small block cluster subsampling take-everys have been set so that the expected number of clusters from each stratum will be an integer.

III. Process

In this process the block clusters will be put into eleven different small block cluster subsampling strata using information from the Sample Design File and the independent listing results. The A.C.E. clusters from the medium, large, and AIR sampling strata are not part of the small block cluster subsampling process, and therefore will be retained in this operation. However, we still assign these clusters to small block cluster subsampling strata and pick up their independent list housing unit (HU) counts. This information is used in large block cluster subsampling. All small clusters that are AIR, AIC, and List/Enumerate as well as those with 10 or more housing units on either the DMAF or IL will be retained in this process.

All the following steps should only be completed for clusters that are in the A.C.E. reduced sample (Current Sample Indicator = 1 on the Sample Design File).

A.   Assigning Clusters to a Small Block Cluster Subsampling Stratum

1.   Using the results of the independent list keying operation, obtain the number of IL HUs for each block cluster and write this value to the variable NHUIL on the Sample Design File. See Attachment B for a layout of the Sample Design File. Include HUs with all eight of the Unit Status Codes[3] in this IL count.

2.   For each block cluster create a new variable, LARGERHU, which is equal to the larger of the DMAF HU count (from the Sample Design File) and IL HU count. For some block clusters such as List/Enumerate, there will be no DMAF count. In these cases assign the IL HU count to the LARGERHU variable. Write this variable to the Sample Design File.

3.   Using the size category (SIZECAT), the American Indian Country Indicator (AICIND), the larger of the DMAF or IL count (LARGERHU)

---

[3]The Unit Status codes are:

1) Occupied or vacant and intended for occupancy	5) Boarded up
2) Under construction	6) Storage of household goods
3) Future construction	7) Vacant mobile home site
4) Unfit for habitation	8) Other

and the List/Enumerate flag (LEIND) from the Sample Design File, assign each cluster a small block cluster subsampling stratum code based on the following table. Write these values to the variable SBCSS on the Sample Design File.

Table 1. Small Block Cluster Subsampling Strata Assignment Rules

IF				THEN
Original Cluster Size Category (SIZECAT)	Larger of DMAF/IL HU count (LARGERHU)	American Indian Country Indicator[4] (AICIND)	List/ Enumerate Indicator (LEIND)	Sub-Sampling Stratum Code (SBCSS)
Small (0-2)	0-2	0 (non AIR/AIC)	0 (not L/E)	01
			1 (L/E)	05
		1 (AIR)	0 or 1	07
		2 (AIC)	0 or 1	08
Small (0-2)	3-5	0 (non AIR/AIC)	0 (not L/E)	02
			1 (L/E)	06
		1 (AIR)	0 or 1	07
		2 (AIC)	0 or 1	09
Small (0-2)	6-9	0 (non AIR/AIC)	0 (not L/E)	03
			1 (L/E)	06
		1 (AIR)	0 or 1	07
		2 (AIC)	0 or 1	09
Small (0-2)	10+	0, 1, or 2	0 or 1	04
Medium (3-79)	all	0, 1, or 2	0 or 1	10
Large (80+)	all	0, 1, or 2	0 or 1	11

[4]For the American Indian Country Indicator :
    0 = Not American Indian Country
    1 = American Indian Reservation or Trustland
    2 = Tribal Jurisdiction statistical area/Alaska Native Village statistical area/tribal designated statistical area

Stratum 01 – Non L/E Small Block Clusters where $0 \leq$ LARGERHU ≤ 2 not on AIR or AIC
Stratum 02 – Non L/E Small Block Clusters where $3 \leq$ LARGERHU ≤ 5 not on AIR or AIC
Stratum 03 – Non L/E Small Block Clusters where $6 \leq$ LARGERHU ≤ 9 not on AIR or AIC
Stratum 04 – Small Block Clusters where $10 \leq$ LARGERHU
Stratum 05 – L/E Small Block Clusters where $0 \leq$ LARGERHU ≤ 2 not on AIR or AIC
Stratum 06 – L/E Small Block Clusters where $3 \leq$ LARGERHU ≤ 9 not on AIR or AIC
Stratum 07 – American Indian Reservation where $0 \leq$ LARGERHU ≤ 9
Stratum 08 – American Indian Country where $0 \leq$ LARGERHU ≤ 2
Stratum 09 – American Indian Country where $3 \leq$ LARGERHU ≤ 9
Stratum 10 – Medium Block Clusters
Stratum 11 – Large Block Clusters

B.   Sending Clusters to Housing Unit Matching

Some of the strata will not be subsampled. Since it is important to start housing unit matching as soon as possible, clusters in these strata should be sent to housing unit matching in a timely manner to prevent any delay in operations.

1.   Create a new variable SB on the Sample Design File to indicate whether a cluster has been retained during the small block cluster subsampling operation. Since all the clusters in small block cluster subsampling strata 04, 07, 10, and 11 will be retained in sample, create and set the random start variable (RSSB) for these cluster to 1.000000, the initial and final take-every (ITESB and FTESB) to 1.000000, and set SB equal to one. Then send them to housing unit matching.[5]

2.   The small clusters in small block cluster strata 01, 02, 03, 05, 06, 08, and 09 will be subsampled to determine which clusters will be sent to HU matching.

C.   Subsample the Small Block Clusters in Strata 01, 02, 03, 05, 06, 08, and 09

For the strata that will be subsampled, the small block cluster subsampling will be done separately for each small block cluster subsampling stratum within each state. The subsampling take-everys will be set so that the expected number of clusters from each small block cluster subsampling stratum is an integer. Subsampling for a specific state should not begin until every small cluster in that state has been assigned to a small block cluster subsampling stratum.

---

[5]It turns out that all clusters in strata 05, 06, 08, and 09 will also be retained in the sample. However, when this specification and the computer programs to implement the operation were being created that was not yet known. We ensure that these clusters are selected by setting their take-everys to one in the small block cluster subsampling parameter file.

1. Sort the block clusters within each small block cluster subsampling stratum (SBCSS) by estimated cluster urbanization (ECLUSURB), county (COUNTY), and A.C.E. cluster number including check digit (CLUST). This sort will help to insure that our sample is representative across these geographic levels.

2. Within each small block cluster subsampling stratum, create an index by numbering the block clusters from 1 to N where N is the number of block clusters in the subsampling stratum.

3. For each small block cluster subsampling stratum, get the initial take-every (ITESB) from the Small Block Subsampling Parameter File that the Sample Design Team has provided. Write this value to the Sample Design File. See Attachment C for a layout.

4. If the number of clusters, N, in a small block cluster subsampling stratum does not equal zero and is less than the initial take-every for that stratum (ITESB) then set the final stratum take-every (FTESB) to the number of clusters in the stratum. Otherwise, if the number of clusters in the small block cluster subsampling stratum is zero or greater than the ITESB set the final stratum take-every (FTESB) equal to the initial stratum take-every (ITESB). This is done to insure that we will select at least one cluster from each small block cluster subsampling stratum.[6]

5. Generate a sequence of numbers L_1, L_2, ..., L_n as follows:

   a. For each subsampling stratum, generate a random number (RN) between 0 and 1 ($0 < RN \leq 1$) with 10 decimal places.

   b. Calculate a random start, RSSB, which equals RN×FTESB. Round this number to six decimal places and write it to the Sample Design File record for each cluster in the subsampling stratum.

   c. Let L_1 = RSSB.

   d. Calculate $L_J = L_{J-1}$ + FTESB for J = 2, 3, ..., n, where n is the largest integer such that [RSSB + (n - 1)×FTESB] $\leq$ N.

---

[6]This step in the specification is no longer necessary because before the take-everys are included in the small block cluster subsampling parameter file they will be computed to ensure that they yield integer expected sample sizes for each subsampling stratum. However, when the specification and computer programs were being written it was not yet known that the take-everys would be computed in this way.

7

e. Round each L_J up to the nearest integer (an integer rounds to itself).

6. For each cluster in the subsampling stratum with an index equal to the rounded values of L_J, $J = 1, 2, ..., n$, assign SB = 1. These block clusters are in the sample. Send them on to housing unit matching.

7. For each cluster in the subsampling stratum with an index not equal to the rounded values of L_J, $J = 1, 2, ..., n$, assign SB = 0. These block clusters are not in the sample. For these clusters set the Current Sample Indicator on the Sample Design File to 0.

8. For each subsampling stratum calculate a check value C such that:

$$C = | (N / FTESB) - n |$$

N = Number of clusters in the subsampling stratum
n = Number of clusters selected from the subsampling stratum
FTESB = Final small block cluster subsampling stratum take-every

If the sampling procedure was performed correctly, then C will be less than one. If C is greater than or equal to one, then contact the author so that operations can be reviewed.[7]

D. Calculate Cluster Weights

For all clusters in the A.C.E. sample after small block cluster subsampling (small, medium, large, and AIR), compute the variable WEIGHTC which is equal to the unbiased weight of each cluster. Calculate this value for each cluster by multiplying the take-everys from the initial block cluster sampling, the take-every from the A.C.E. reduction and the final take-every from small block cluster subsampling. Round to six decimal places and write to the Sample Design File. For all clusters not in the A.C.E. sample after small block cluster subsampling leave this value blank.

$$WEIGHTC = TE1 \times TE2 \times TEAR \times FTESB$$

---

[7]Before the take-everys were written to the small block clusters subsampling parameter file they were computed to yield integer expected sample sizes from all the subsampling strata. Because of this, C will always be equal to 0. This was not written into the specification because at the time of its writing we did not know the take-everys would be computeded in this way.

8

E.   Produce Verification Output

   1.   Provide the sampling staff with access to the Cluster Status File, so that with the Sample Design File we may replicate the sampling operation.

   2.   Create the Block Cluster Subsampling Verification File. This file provides summary information for the different strata in each state and will be used by the sampling staff for verification. See Attachment D for a layout. Several calculations are required for this file.

      a.   To calculate the average LARGERHU for all clusters in the stratum add up the LARGERHU values in the stratum and divide by the number of clusters in the stratum.

      b.   To calculate the average LARGERHU of clusters selected from a stratum add up the LARGERHU values of the clusters selected from the stratum and divide by the number of clusters selected from the stratum.

      c.   For small block cluster subsampling strata one through nine, calculate the weight of a cluster selected from a stratum by multiplying TE1 × TE2 × TEAR × FTESB. For strata one through nine the weights of each cluster within a stratum within a state should be equal.

   3.   Create the Independent List Housing Unit Information File. Please provide us this file as soon as possible so that we may set our final take-everys. This file is required in the setting of the take-everys because we want to insure that the expected number of clusters from each small block cluster subsampling stratum is an integer. This file will also be used by the sampling staff for evaluation purposes. See Attachment E for a layout.

IV. Input

The following files are inputs to this operation.

A.   Cluster Status File

This file contains one record for each block cluster selected in the A.C.E. listing sample. The original source of this file is the Sample Design File. It is updated with information from other processing stages. For small block cluster subsampling, this file is used to obtain the keyed and valid counts of Independent Listing housing units by type of unit status. For information about this file contact Courtney Ford in the Processing Support & A.C.E. Systems Staff at 301-457-4121.

B.   Sample Design File, Version 2

This file contains information about the entire sampling history of each block cluster. See Attachment B for a layout. Version 2 reflects listing sample selection and A.C.E. cluster reduction. There are a total of 29,717 cluster records on the file. Only clusters with CSI= 1 are in the reduced A.C.E. sample. Note that once a cluster drops out of sample, the fields from the remaining operations will be left blank.

C.   Small Block Cluster Subsampling Parameter File

This file, which contains one record for each state, records the initial take-every for each of the eleven small block cluster subsampling strata. These initial take-everys will be set so that the expected sample size from each small block cluster subsampling strata will be an integer. See Attachment C for a layout. This file will be created in late January by the Sample Design Team.

V. Output

A.   Sample Design File, Version 3

This file contains information about the entire sampling history of each block cluster. See Attachment B for a layout. It will be updated after the small block cluster subsampling process.

B.   Small Block Cluster Subsampling Verification File

This file will be used to assist the sample design staff in verification procedures. See Attachment D for a layout.

C.   Independent Listing Housing Unit Information File

This file will be used by the sample design staff to evaluate the stratification of the small block clusters. It will contain one record for each cluster (including medium, large, and AIR clusters) still in sample before Small Block Cluster Subsampling. See Attachment E for a layout.

cc: DSSD Census 2000 Procedures and Operations Memorandum Series Distribution List
A.C.E. Team Leaders
Statistical Design Team Leaders
Sample Design Team
Estimation Team
Variance Estimation Team

Approximate Small Block Cluster Subsampling Take-Everys

Listed below are the approximate small block cluster subsampling take-everys. These take-everys have not been computed to insure that the expected number of clusters from each stratum is an integer. This computation takes place before the take-everys are written to the small block cluster subsampling parameter file, and it is not part of this specification. The final take-everys, as well as the exact methodology for their calculation, are forthcoming in a later document. Because the final take-everys will be different than the ones listed below, this attachment is included merely as a guide.

FIPS CODE	State Name	Stratum 1	Stratum 2	Stratum 3	Stratum 4-9
01	Alabama	10.00	4.00	2.22	1.00
02	Alaska	10.00	4.00	2.22	1.00
04	Arizona	2.80	1.12	1.00	1.00
05	Arkansas	8.97	3.59	1.99	1.00
06	California	2.57	1.03	1.00	1.00
08	Colorado	4.11	1.64	1.00	1.00
09	Connecticut	8.71	3.48	1.93	1.00
10	Delaware	10.00	4.00	2.22	1.00
11	District of Columbia	10.00	4.00	2.22	1.00
12	Florida	4.66	1.86	1.04	1.00
13	Georgia	10.00	4.00	2.22	1.00
15	Hawaii	3.92	1.57	1.00	1.00
16	Idaho	2.09	1.00	1.00	1.00
17	Illinois	10.00	4.00	2.22	1.00
18	Indiana	10.00	4.00	2.22	1.00
19	Iowa	10.00	4.00	2.22	1.00
20	Kansas	10.00	4.00	2.22	1.00
21	Kentucky	10.00	4.00	2.22	1.00
22	Louisiana	3.59	1.43	1.00	1.00
23	Maine	6.24	2.50	1.38	1.00
24	Maryland	9.48	3.79	2.11	1.00
25	Massachusetts	8.47	3.39	1.88	1.00
26	Michigan	9.22	3.69	2.05	1.00
27	Minnesota	10.00	4.00	2.22	1.00
28	Mississippi	6.80	2.72	1.51	1.00
29	Missouri	10.00	4.00	2.22	1.00
30	Montana	2.68	1.07	1.00	1.00
31	Nebraska	10.00	4.00	2.22	1.00
32	Nevada	1.90	1.00	1.00	1.00
33	New Hampshire	10.00	4.00	2.22	1.00
34	New Jersey	5.51	2.20	1.22	1.00
35	New Mexico	2.16	1.00	1.00	1.00

36	New York	9.94	3.98	2.21	1.00
37	North Carolina	10.00	4.00	2.22	1.00
38	North Dakota	8.10	3.24	1.80	1.00
39	Ohio	10.00	4.00	2.22	1.00
40	Oklahoma	9.70	3.88	2.16	1.00
41	Oregon	2.11	1.00	1.00	1.00
42	Pennsylvania	10.00	4.00	2.22	1.00
44	Rhode Island	10.00	4.00	2.22	1.00
45	South Carolina	10.00	4.00	2.22	1.00
46	South Dakota	8.63	3.45	1.92	1.00
47	Tennessee	10.00	4.00	2.22	1.00
48	Texas	3.20	1.28	1.00	1.00
49	Utah	2.17	1.00	1.00	1.00
50	Vermont	10.00	4.00	2.22	1.00
51	Virginia	9.03	3.61	2.01	1.00
53	Washington	3.00	1.20	1.00	1.00
54	West Virginia	7.05	2.82	1.57	1.00
55	Wisconsin	10.00	4.00	2.22	1.00
56	Wyoming	1.94	1.00	1.00	1.00
72	Puerto Rico	3.31	1.32	1.00	1.00

Sample Design File Layout

The Sample Design File contains one record per block cluster selected during the initial block cluster sampling. If the block cluster is subsampled out of sample during the second step of sampling, the A.C.E. reduction or during small block subsampling, the remaining variables will be left blank. The initial version of the file, which will be created following the initial block cluster selection, is called SDF.US1. For each subsequent update to the file, the version number will increase by one (i.e. SDF.US2, SDF.US3). The layout for the Sample Design File is as follows:

Variable Description	Name	Places	Source
Census Region	REGION	1	UN
Census Division	DIV	2	UN
State code	STATE	3-4	UN
County code	COUNTY	5-7	UN
Local census office	LCO	8-11	CS
Interim Tract (Pseudo Tract)	ITRACT	12-17	BC
Current Sample Indicator	CSI	19	UO
A.C.E. block cluster number	CLUST	21-25	CS
Check Digit	DIGIT	26	CS
Geography block cluster number	GCLUST	28-32	BC
List/Enumerate Indicator (1= L/E, 0 = Non-L/E)	LEIND	33	BC
Type of Enumeration Area Recode	TEACR	34	CS
Type of Enumeration Area group	TEAG	36	BC
Number of HUs used for sample design	NHU	37-41	BC
Number of MAF HUs	NHUM	43-47	BC
Number of 1990 HUs	NHU90	49-53	BC
Sampling Stratum	SS	55	UN
1 = Small			
2 = Medium			
3 = Large			
4 = American Indian Reservation			
American Indian Country Indicator	AICIND	56	BC
0 = No American Indian Country			
1 = American Indian Reservation/trust land			
2 = Tribal Jurisdiction Statistical Area/			
Tribal Designated Statistical Area/			
Alaska Native Village Statistical Area			
Demographic/Tenure Group code	DTCODE	57-58	UN
Demographic/Tenure Group label	DTLABEL	59-60	UN
Estimated Urbanicity of block cluster	ECLUSURB	62	UN
1 = Urban Area with population ≥250,000			
2 = Other Urban Area			
3 = Non-Urban Area			

Variable Description	Name	Places	Source
Size Category	SIZCAT	63	UN
1=Small (0-2 hus)			
2=Medium (3-79 hus)			
3=Large (80+ hus)			
Additional space		64-91	

Variable Description	Name	Places	Source
First step index number	INDEX1	92-99	CS
Listing sample selection indicator	BC1	101	CS
1 = Selected			
Random start for listing sample selection	RS1	103-113	UN
Take-every for listing sample selection	TE1	115-125	UN
Second step listing sample selection indicator	BC2	127	CS
0 = Not Selected, 1 = Selected			
Random start for second step listing sample selection	RS2	129-139	CS
Take-every for second step listing sample selection	TE2	141-151	CS
Unbiased weight after listing sample selection	WEIGHTBC	153-164	CS
Additional space		165-175	

Variable Description	Name	Places	Source
Preliminary Number of HUs on the Independent List	NHUILP	176-180	AR
Number of Housing Units On January 2000 DMAF	NHUDMAF	182-186	AR
Demographic code	DEMCODE	188-188	AR
1 = Minority			
2 = Non-minority			
3 = Puerto Rico			
Consistency Code	CONCODE	189-189	AR
1 = Low Inconsistent (IL significantly smaller than DMAF)			
2 = Consistent			
3 = High Inconsistent (IL significantly larger than DMAF)			
A.C.E. reduction stratum	ARS	190-191	AR
A.C.E. reduction indicator	ACERED	193-193	AR
0 = Not Selected, 1 = Selected			
Random start for A.C.E. reduction	RSAR	195-205	AR
Take-every for A.C.E. reduction	TEAR	207-217	AR
Unbiased weight after A.C.E. reduction	WEIGHTAR	219-230	AR
Collapsing flag	COLFLAG	232-232	AR
A.C.E. Reduction index number	INDEXR	234-241	AR
Number of Housing Units on the December 1999 DMAF (Initial)	NHUDMAFI	243-247	AR
Additional space		248-300	

Variable Description	Name	Places	Source
Number of HUs on the Independent List	NHUIL	301-305	SB
Small Block Cluster Subsampling Stratum	SBCSS	306-307	SB
Small Block Subsampling Indicator	SB	308	SB
0 = Not Selected, 1 = Selected			
Random Start for Small Block subsampling	RSSB	310-320	SB
Initial take-every for Small Block subsampling	ITESB	322-332	SB
Unbiased weight for ACE cluster	WEIGHTC	334-345	SB
Larger of the DMAF and IL HU count	LARGERHU	347-351	SB
Final take-every for Small Block subsampling	FTESB	352-362	SB
Additional space		363-370	SB

Source Codes

AR: ACE Reduction
BC: Block Clustering
CS: Block Cluster Sampling
SB: Small Block Subsampling
UN: Universe File Creation
UO: Updated for each operation

Small Block Cluster Subsampling Parameter File Layout

This file, which will be created by the sample design staff, will provide the take-every for the different strata in each state. There will be one record for each state, the District of Columbia, and Puerto Rico. It will be called SBCSPF.DAT.

Variable Description	Name	Places
State (FIPS Code)	STATE	1-2
Take-every for stratum 01*	TE_1	4-14
Take-every for stratum 02*	TE_2	16-26
Take-every for stratum 03*	TE_3	28-38
Take-every for stratum 04*	TE_4	40-50
Take-every for stratum 05*	TE_5	52-62
Take-every for stratum 06*	TE_6	64-74
Take-every for stratum 07*	TE_7	76-86
Take-every for stratum 08*	TE_8	88-98
Take-every for stratum 09*	TE_9	100-110
Take-every for stratum 10*	TE_10	112-122
Take-every for stratum 11*	TE-11	124-134

*Note: Take-everys will be rounded to six decimal places and may be non-integer values.

Small Block Cluster Subsampling Verification File Layout

This file will be created during processing to assist the sample design staff in verification. One file will be created for the entire nation (including the District of Columbia and Puerto Rico). This file will be called SBCSVF.DAT.

Variable Description	Name	Places
State (FIPS Code)	STATE	1-2
Initial take-every for stratum 01	ITE_1	4-14
Number of clusters in stratum 01	CLUS_1	15-17
Final take-every for stratum 01	FTE_1	18-28
Number of clusters selected from stratum 01	SCLUS_1	29-31
Number of IL HUs in stratum 01	ILHU_1	32-36
Number of IL HUs in clusters selected from stratum 01	SILHU_1	37-41
Average LARGERHU of clusters in stratum 01	ALHU_1	42-47
Rounded to three decimal places		
Average LARGERHU of clusters selected from stratum 01	SALHU_1	48-53
Rounded to three decimal places		
Weight of clusters selected from stratum 01	WEIGHT_1	54-65
Random number used to sample stratum 01	RN_1	66-76
Initial take-every for stratum 02	ITE_2	88-98
Number of clusters in stratum 02	CLUS_2	99-101
Final take-every for stratum 02	FTE_2	102-112
Number of clusters selected from stratum 02	SCLUS_2	113-115
Number of IL HUs in stratum 02	ILHU_2	116-120
Number of IL HUs in clusters selected from stratum 02	SILHU_2	121-125
Average LARGERHU of clusters in stratum 02	ALHU_2	126-131
Rounded to three decimal places		
Average LARGERHU of clusters selected from stratum 02	SALHU_2	132-137
Rounded to three decimal places		
Weight of clusters selected from stratum 02	WEIGHT_2	138-149
Random number used to sample stratum 02	RN_2	150-160
Initial take-every for stratum 03	ITE_3	182-192
Number of clusters in stratum 03	CLUS_3	193-195
Final take-every for stratum 03	FTE_3	196-206
Number of clusters selected from stratum 03	SCLUS_3	207-209
Number of IL HUs in stratum 03	ILHU_3	210-214
Number of IL HUs in clusters selected from stratum 03	SILHU_3	215-219
Average LARGERHU of clusters in stratum 03	ALHU_3	220-225
Rounded to three decimal places		
Average LARGERHU of clusters selected from stratum 03	SALHU_3	226-231
Rounded to three decimal places		
Weight of clusters selected from stratum 03	WEIGHT_3	232-243
Random number used to sample stratum 03	RN_3	244-254

Initial take-every for stratum 04	ITE_4	286-296
Number of clusters in stratum 04	CLUS_4	297-299
Final take-every for stratum 04	FTE_4	300-310
Number of clusters selected from stratum 04	SCLUS_4	311-313
Number of IL HUs in stratum 04	ILHU_4	314-318
Number of IL HUs in clusters selected from stratum 04	SILHU_4	319-323
Average LARGERHU of clusters in stratum 04	ALHU_4	324-329
Rounded to three decimal places		
Average LARGERHU of clusters selected from stratum 04	SALHU_4	330-335
Rounded to three decimal places		
Weight of clusters selected from stratum 04	WEIGHT_4	336-347
Initial take-every for stratum 05	ITE_5	390-400
Number of clusters in stratum 05	CLUS_5	401-403
Final take-every for stratum 05	FTE_5	404-414
Number of clusters selected from stratum 05	SCLUS_5	415-417
Number of IL HUs in stratum 05	ILHU_5	418-422
Number of IL HUs in clusters selected from stratum 05	SILHU_5	423-427
Average LARGERHU of clusters in stratum 05	ALHU_5	428-433
Rounded to three decimal places		
Average LARGERHU of clusters selected from stratum 05	SALHU_5	434-439
Rounded to three decimal places		
Weight of clusters selected from stratum 05	WEIGHT_5	440-451
Random number used to sample stratum 05	RN_5	452-462
Initial take-every for stratum 06	ITE_6	494-504
Number of clusters in stratum 06	CLUS_6	505-507
Final take-every for stratum 06	FTE_6	508-518
Number of clusters selected from stratum 06	SCLUS_6	519-521
Number of IL HUs in stratum 06	ILHU_6	522-526
Number of IL HUs in clusters selected from stratum 06	SILHU_6	527-531
Average LARGERHU of clusters in stratum 06	ALHU_6	532-537
Rounded to three decimal places		
Average LARGERHU of clusters selected from stratum 06	SALHU_6	538-543
Rounded to three decimal places		
Weight of clusters selected from stratum 06	WEIGHT_6	544-555
Random number used to sample stratum 06	RN_6	556-566

Initial take-every for stratum 07	ITE_7	598-608
Number of clusters in stratum 07	CLUS_7	609-611
Final take-every for stratum 07	FTE_7	612-622
Number of clusters selected from stratum 07	SCLUS_7	623-625
Number of IL HUs in stratum 07	ILHU_7	626-630
Number of IL HUs in clusters selected from stratum 07	SILHU_7	631-635
Average LARGERHU of clusters in stratum 07	ALHU_7	636-641
Rounded to three decimal places		
Average LARGERHU of clusters selected from stratum 07	SALHU_7	642-647
Rounded to three decimal places		
Weight of clusters selected from stratum 07	WEIGHT_7	648-659
Initial take-every for stratum 08	ITE_8	702-712
Number of clusters in stratum 08	CLUS_8	713-715
Final take-every for stratum 08	FTE_8	716-726
Number of clusters selected from stratum 08	SCLUS_8	727-729
Number of IL HUs in stratum 08	ILHU_8	730-734
Number of IL HUs in clusters selected from stratum 08	SILHU_8	735-739
Average LARGERHU of clusters in stratum 08	ALHU_8	740-745
Rounded to three decimal places		
Average LARGERHU of clusters selected from stratum 08	SALHU_8	746-751
Rounded to three decimal places		
Weight of clusters selected from stratum 08	WEIGHT_8	752-763
Random number used to sample stratum 08	RN_8	764-774
Initial take-every for stratum 09	ITE_9	806-816
Number of clusters in stratum 09	CLUS_9	817-819
Final take-every for stratum 09	FTE_9	820-830
Number of clusters selected from stratum 09	SCLUS_9	831-833
Number of IL HUs in stratum 09	ILHU_9	834-838
Number of IL HUs in clusters selected from stratum 09	SILHU_9	839-843
Average LARGERHU of clusters in stratum 09	ALHU_9	844-849
Rounded to three decimal places		
Average LARGERHU of clusters selected from stratum 09	SALHU_9	850-855
Rounded to three decimal places		
Weight of clusters selected from stratum 09	WEIGHT_9	856-867
Random number used to sample stratum 09	RN_9	868-878

Initial take-every for stratum 10	ITE_10	910-920
Number of clusters in stratum 10	CLUS_10	921-923
Final take-every for stratum 10	FTE_10	924-934
Number of clusters selected from stratum 10	SCLUS_10	935-937
Number of IL HUs in stratum 10	ILHU_10	938-942
Number of IL HUs in clusters selected from stratum 10	SILHU_10	943-947
Average LARGERHU of clusters in stratum 10 *Rounded to three decimal places*	ALHU_10	948-953
Average LARGERHU of clusters selected from stratum 10 *Rounded to three decimal places*	SALHU_10	954-959
Initial take-every for stratum 11	ITE_11	1014-1024
Number of clusters in stratum 11	CLUS_11	1025-1027
Final take-every for stratum 11	FTE_11	1028-1038
Number of clusters selected from stratum 11	SCLUS_11	1039-1041
Number of IL HUs in stratum 11	ILHU_11	1042-1046
Number of IL HUs in clusters selected from stratum 11	SILHU_11	1047-1051
Average LARGERHU of clusters in stratum 11 *Rounded to three decimal places*	ALHU_11	1052-1058
Average LARGERHU of clusters selected from stratum 11 *Rounded to three decimal places*	SALHU_11	1059-1065

Independent List Housing Unit Information File Layout

This cluster level file will be created during processing to assist the sample design staff in evaluation of the subsampling. It may also be used to help verify the large block cluster subsampling parameters. One file will be created for the nation (including the District of Columbia and Puerto Rico). The file will include one record for each cluster (including medium, large, and AIR clusters) in sample before Small Block Cluster Subsampling. The counts on this file will be keyed and valid IL counts. The file will be called ILHUIF.DAT.

Variable Description	Name	Places
State (FIPS Code)	STATE	1-2
County	COUNTY	3-5
A.C.E. Cluster Number	CLUST	6-10
Check Digit	DIGIT	11-11
Total number of IL HUs	ILHU	12-16
Number of HUs where USTAT = 1	USTAT_1	17-21
(Occupied or vacant and intended for occupancy)		
Number of HUs where USTAT = 2	USTAT_2	22-26
(Under construction)		
Number of HUs where USTAT = 3	USTAT_3	27-31
(Future Construction)		
Number of HUs where USTAT = 4	USTAT_4	32-36
(Unfit for Habitation)		
Number of HUs where USTAT = 5	USTAT_5	37-41
(Boarded Up)		
Number of HUs where USTAT = 6	USTAT_6	42-46
(Storage of household goods)		
Number of HUs where USTAT = 7	USTAT_7	47-51
(Vacant mobile home site)		
Number of HUs where USTAT = 8	USTAT_8	52-56
(Other)		